

Uttering sentences made up of words and gestures¹

Philippe De Brabanter

Institut Jean Nicod – Université Paris 4-Sorbonne

1. Setting the scene

In several places, Robyn Carston warns the student of utterance interpretation against neglect of the “fact that most verbal utterances are a complex of linguistic, paralinguistic, facial and vocal gestures, which appear to function as a signal receiving a unified interpretation”, a fact which, she writes, remains “rather under-explored” (2004, p. 824). Such neglect is unjustified because, she writes, “the domain of pragmatics is a natural class of environmental phenomena, that of ostensive (=communicative) stimuli; verbal utterances are the central case, but not the only one, and they themselves are frequently accompanied by other ostensive gestures of the face, hands, voice etc, all of which have to be interpreted together if one is to correctly infer what is being communicated” (2002, p. 129; see also Carston & Powell, p. 16 of online version). This position rests on the assumption that there is a single “pragmatic system” at work in the interpretation of “ostensive stimuli”. When it comes to interpreting verbal stimuli, the same mechanisms and resources are used as when it comes to processing non-verbal ones. If there is no distinct “linguistic pragmatic system”, then the scholar who studies communication should not favour the verbal at the expense of the non-verbal.

In a more radical vein, Herb Clark (1996 and 1997, *passim*) contends that theories of “language use” – into which he includes most pragmatics originating in Grice – are essentially misguided in their almost exclusive focus on linguistic *symbols* (conventional signs) and neglect of *indices* (signs causally related to their object) and *icons* (signs related to their object in virtue of some shared quality), to use C.S. Peirce’s terminology. Ignorance of the prevalence of *composite signals* makes standard pictures of utterance interpretation fundamentally incomplete: “Ignoring nonlinguistic methods has distorted people’s picture of language use, and it is important to put that picture right” (1996, p. 156).

One may have reservations about Clark’s use of the term “language use” and Carston has pointed out that Clark’s terminological choice is potentially misleading as it refers to phenomena that “include many non-linguistic behaviours and exclude many (non-communicative) employments of linguistic forms” (1999, 167). But, as the quotations in the first paragraph testify, Carston (and Relevance theory more generally) agrees with Clark that human communication is usually multi-modal.

In spite of this, which I take to be an empirical observation, it is clear that semanticists but also pragmaticists usually give verbal aspects of communication preferential treatment. As a result, the non-verbal aspects of an act of communication are cut off from the verbal ones and end up being treated *as if* they were not underlain by communicative intentions. They are in fact relegated to the context of utterance, on a par with other aspects of the context that can be useful in interpreting the utterance (location, time, perceptible objects and events, shared background, etc.). It is true that, once an ostensive stimulus has been processed, it becomes part of the context against which a subsequent act of communication will be interpreted. But in that respect gesturing is no different from the use of words: it is only after it has been processed *qua* communicative stimulus that it can be treated as an element of the context of

¹ Very special thanks to Neftali Villanueva Fernandez and Mikhail Kissine for their insightful and helpful remarks on a draft of this paper. Thanks also to Steven Davis, Esther Romero, Belén Soria, Dan Sperber, and benevolent audiences in Granada in 2004 and 2006 for various discussions and comments on the ideas presented in this paper. Finally, I wish to express my gratitude to Peter Elliott for supplying the illustrations.

utterance.² That communicative role of non-verbal stimuli, especially icons (cf. Clark, 1996, p. 188) is often passed over in the semantics and pragmatics literature. In this paper, I try to make up for this situation, though on a very limited scale: I focus on the interpretation of a variety of multi-modal signals comprising conventional signs (words) and icons.

2. The data

My data consist of a special subset of multi-modal signals, namely ostensive stimuli that mix the verbal with the non-verbal in such a way that a gesture, posture, facial expression (or a combination thereof) seems to stand in for a linguistic constituent which remains unrealised. Here is a real-life example, translated from French (the speaker is a fretting but relieved regular customer addressing an assistant in a fashion shop):

- (1) I didn't see the [IMITATION OF FRIGHTENING GRUMPINESS] woman today; will she be back this week?

The square-bracketed string in small capitals is an interpretation of the facial expressions and gestures performed in the conversational setting. This imitation of grumpiness is a *demonstration* in Clark & Gerrig's sense (1990), that is, an act of "illustration by exemplification".³ In (1), that demonstration plays a similar role to an Adjective Phrase. It is a more entertaining, livelier way of conveying the same piece of information as could have been communicated by uttering "frighteningly grumpy" or some such phrase. The argument I develop in later sections is that the non-verbal stimulus in (1) is embedded into a broader linguistic structure – an NP⁴ and, ultimately, a sentence – and that it performs a genuine linguistic function.

The relevant data for this paper have to be distinguished from two related sets of examples: (i) spoken utterances of complete sentences which, as is usual, are accompanied by some gesturing; (ii) non-verbal communicative stimuli that do not appear to stand in for an unrealised linguistic constituent, either because they do not co-occur with words, or because the words they co-occur with do not add up to a sentential structure.

I want to leave the (i) cases aside. Not because I assume that the gesturing usually accompanying speech does not contribute to communicated content. It most probably does in at least some cases (cf. Wilson, 2000, p. 431f; see also section 3). But this gesturing does not play a *linguistic* part; it does not fulfil a syntactic function.

As for the (ii) cases, there are many illustrations in the literature. Take Sperber & Wilson's example of the person sniffing ostensively to signal the presence of gas (1995, p. 55), a situation in which no words are uttered.⁵ Or take this other example from Clark and Gerrig, in which a little boy draws the attention of his adult friend:

- (2) Herb! [points to Eve] + [puts an imaginary camera to his eyes and clicks the shutter].
(1990, p. 765)

² I give a somewhat more detailed discussion of this issue in De Brabanter (2005). Note that I will adopt relevance-theoretic parlance and talk of *ostensive stimulus* and *communicative stimulus*. As far as I can see, these terms are interchangeable with Herb Clark's *signal*.

³ In the rest of this article, I will use "demonstration" in this iconic sense. Care should be taken not to confuse it with the more widespread use of "demonstration" in philosophy of language to designate an act of using a demonstrative (i.e. an indexical sign).

⁴ I shall use this symbol for what many would gladly call a *Determiner Phrase* or *DP*. Nothing hinges on this terminological choice.

⁵ An excellent discussion of instances of entirely non-verbal communication can be found in Sperber & Wilson (1995, pp. 48-56).

The child means to inform Herb that Eve is taking a photograph. This he achieves by producing various types of stimuli, linguistic and non-linguistic. The second (the manipulation of an imaginary camera) is a demonstration, whereas the first is an *indication* (an indexical gesture). Here, there is no embedding of the gesturing into a linguistic structure, as was arguably the case in (1). Hence, there is no point in analysing the gestures in (2) in linguistic terms, as opposed to the analysis I advocate with respect to (1).

3. Acts of ostensive communication?

My analysis starts from the assumption that (certain) co-verbal gestures are ostensive stimuli. Thus, I believe the demonstrations and indications represented in (1) and (2) to be underlain by a communicative intention, just as much as the verbal part of the signals. Some caution is in order, however. For instance, Clark has an example in which placing a candy wrapper in a litter basket is said to *indicate* that it is waste (2003, p. 257). Surely this can be done conspicuously, with a communicative intention, but it need not be: most of the time when I dispose of a candy wrapper, I do not intend to communicate that I am treating it as waste. Therefore, when Clark makes the sanguine claim that “[m]ost iconic gestures are genuine signals by which speakers mean things” (1996, p. 177), we are well-advised to take this with a pinch of salt.

In a study whose main goal is to show that at least *some* iconic gestures are communicative stimuli, Melinger & Levelt (2004) devote some time to reviewing the literature on gestures and communicative intentions. They point out that, whereas it is widely agreed that “co-verbal” indexical gestures are often communicative,⁶ “few studies have explicitly addressed whether iconic gestures form part of the speaker’s communicative intention” (2004, p. 121). Apparently, many prior studies exhibited flaws in their methodology – the few examples discussed by Melinger & Levelt indeed prove to be less than perfect. Their conclusion is that, up until their own paper, it was not at all well-established that co-verbal iconic gestures are (sometimes) communicative.

So, what of the iconic gestures that occur in my data? How can it be shown that their communicative status is above suspicion? My chief argument is that the demonstrations involved are what Clark (1996, p. 178) calls *component gestures*, as opposed to *concurrent* ones. More precisely, they fall into one of the three categories of component gestures distinguished by Clark, namely those cases in which all there is (at a certain point in a signal) is the iconic gesture. In the examples examined in sections 7 and 8 below (examples 10, 11, 15, 16 and 17), the demonstrations clearly fill “informative gaps”; they are part and parcel of the explicit content of the utterance. If they are not processed, the speaker’s informative intention is not satisfied. This proves less evident in (1), because the demonstration occurs as part of a referential NP that contributes an individual rather than a description to the proposition expressed. Thus, the content of the demonstration does not end up in the propositional content. However, the demonstration helps (and is, I believe, *meant* to help) the shop assistant figure out the referent of the definite description “a ... woman”. I have to admit, still, that this may fall short of establishing the communicative status of the demonstration in (1).

The final bit of evidence I wish to adduce in support of the communicative status of the demonstrations in my data is based on the fact that, if the gesturing is not interpreted as being communicative, it will in most cases feel extremely odd. Take (1) again: absent a

⁶ In my remarks about the candy wrapper I do not wish to say that indications are not underlain by communicative intentions. Rather, I believe that Clark is wrong to treat this example as clearly involving an indication.

communicative intention, the mimicry would have to be seen as some reflex contraction of the speaker's face and would therefore feel like a sudden incomprehensible intrusion if it was not grasped as depicting a property of "the woman". In the examples examined in section 7 and 8, the situation is the same: if not recognised as contributing to the communicated content, the gesturing would appear as a sudden voluntary and rude disruption of communication, whereas in reality communication continues smoothly as the speaker goes from words to non-verbal stimuli.

4. Non-sentential speech acts?

The discussion so far falls short of establishing that example (1) is well and truly a sentence, though one made up of verbal and non-verbal elements. Before bringing further arguments in favour of that claim, I want to get one alternative account out of the way. For some time now, there has been a heated debate within the semantics/pragmatics community as to whether there are such things as non-sentential (or sub-sentential) speech acts. The stakes are high: their existence would threaten a popular version of how the compositionality thesis can be articulated with semantic theory. On this view, as expounded f.i. in Stanley (2000), "all effects of extra-linguistic context on the truth-conditions of assertions are traceable to logical form" and "composition rules do not vary as a function of extra-linguistic context" (Stanley, 2000, p. 392 and 395). If non-sentential speech (esp. assertions) exists, either some truth-conditional determinants are *not* found in logical form or some composition rules *do* vary across contexts.⁷

In the debate about non-sentential speech, two main positions can be distinguished: on the one hand, there are those like Stainton, Elugardo and Robyn Carston who hold that "[s]peakers can make assertions by uttering ordinary, unembedded, words or phrases" (Stainton 1995, p. 281).⁸ Here is an illustration offered by Carston.

it's breakfast time and, coming into the kitchen, I see my companion searching around in the lower reaches of a cupboard; knowing his breakfast habits, I guess that he's looking for a jar of marmalade and I utter:
[(3)] On the top shelf. (2002, p. 130)

On a Stainton-like analysis, this utterance of (3) is a bona fide speech act: it has a rather definite propositional content and an assertive illocutionary force. If someone were to ask what the utterer has said, one could answer that she has said that the marmalade is on the top shelf. In other words, it is an assertion that has truth-conditions and that can be reported indirectly. If the jar turned out not to be on the top shelf, the speaker would be judged to have spoken falsely.

In this case, opponents of Stainton's position, like Jason Stanley, Peter Ludlow or Jason Merchant would agree that (3) is uttered to perform a genuine illocutionary act. But they would argue that (3) does not show its whole syntactic structure. In other words, they would argue that (3) is an instance of syntactic ellipsis.⁹

⁷ There exist other conceptions of compositionality, but the one just portrayed is a popular one and any serious challenge to it is a matter of importance to the semantic community at large.

⁸ The debate gives pride of place to *assertions*. It may be that some writers do not reject non-sentential speech in general, only non-sentential assertions. Still, someone like Jason Stanley assumes that his objections to non-sentential assertions "generalize to other speech acts" (2000, p. 392), and so does Peter Ludlow (2005).

⁹ With respect to some of Stainton's alleged instances of non-sentential speech, Stanley adopts a different strategy: he argues that they are simply not "linguistic speech acts".

I am not certain who is right: is (3) a genuine non-sentential assertion? Mikhail Kissine (p.c.) tells me that a similar phrase in Russian would occur in the ablative, and that this ablative would be assigned by a “location” verb (in other cases, e.g. with verbs like “put”, the same phrase would occur in the accusative)¹⁰. If it turns out that case must be assigned by a hierarchically superior constituent, then one obvious way of accounting for (3) is to say that it involves syntactic ellipsis (of a case-assigning verb). However, this debate is a very complex one and it goes well beyond the objectives of this paper. For the time being, my main concern is to show that Stainton’s analysis – which I believe to be correct in *some* cases – cannot apply to the data I am concerned with here.

First, note that, in the situation described, (3) is not accompanied by any particular non-verbal communicative stimuli. It is self-sufficient as an utterance. Granted, for (3) to make sense some assumptions must be mutually manifest to speaker and hearer (e.g. that the hearer wants marmalade for breakfast). In particular, there must be an object in the context to which the property literally expressed by (3) can be anchored. Although this may make a difference to an opponent of non-sentential speech, who may find something here to support the ellipsis account, it makes none to me: what matters is that (3) requires no demonstration or indication, i.e. no communicative gesturing, to be fully understood. This, I believe makes it substantially different from my example (1). Second, from a syntactic point of view, Stainton (2005, p. 384) argues that examples like (3) have the structure of a lexically headed phrase (the head here is Prep) rather than of an inflectional phrase, i.e. in traditional terms a clause. This, once again, is different from (1), where the presence of the verb “didn’t see” means that we have a clausal structure headed by an inflectional element. Most of the examples I look at in sections 7 and 8 are like (1) in this respect.

I therefore conclude that Stainton’s analysis does not apply to the core data of this paper. Yet, there is a potential lesson from Stainton’s study of non-sentential speech that I wish to highlight. If it is correct in at least some cases, his analysis pre-empts two tempting arguments for anyone trying to establish that utterances like (1) are genuine sentences. Neither the fact that (1) has truth-conditions nor the fact that it can be reported using indirect speech is proof that it is a sentence: utterances of structures other than sentences exhibit these characteristics too.

5. Quotations and demonstrations

I now turn to an outline of the theory that should enable me to establish that utterances like (1) can be said to be genuine sentences. I will start by showing that it is legitimate to analyse the sort of non-verbal “intrusion” occurring in (1) in the same way we analyse quotations. The central idea is that quotations are a variety of demonstrations (Clark & Gerrig, 1990), an idea revived and fleshed out by Recanati not long ago (2000, 2001). Direct speech offers a good illustration of what it means to treat quotations as demonstrations. When I quote – say Keith’s words – I can obviously not produce those very tokens that were uttered by Keith: they are one-time occurrences. What I do instead is produce an utterance-token that instantiates the same utterance-type as the original words. In other words, I exemplify that type. Take:

(4) Keith said, “I think tha’’s all a pile of roobish”.

In (4), the words in quotation marks are displayed; i.e. the addressee’s attention is drawn to them, but only inasmuch as they exemplify (the same type as) Keith’s earlier utterance.

¹⁰ A similar situation obtains in German, where the NP that follows “auf” (= “on”) can be assigned either the accusative with a destination verb or the dative with a location verb. In this case, the German translation of “on the top shelf” would occur in the dative.

Actually, the displayed token in (4) is likely to instantiate several different types or clusters of properties,¹¹ only one of which is the *utterance*-type. Among the instantiated types, only some are targets of the depiction, and these depicted targets need not even include the utterance-type. Assuming that I have uttered (4) aloud, my deviant pronunciation of “rubbish” and omission of a final “-t” may indicate that I intended to demonstrate other properties exemplified by the words in quotes. For instance, I may have wished to demonstrate Keith’s way of talking (his accent), or more broadly, the pronunciation characteristic of the dialect spoken by Keith, say working-class northern English. Or perhaps I wanted to demonstrate Keith’s bad manners, or what have you. It is the target the quoter has in mind which determines which aspects of a demonstration are “depictive”. But, in the end, all it takes for a successful demonstration is that, in the context of utterance, some link be recoverable between the quoted words and the property demonstrated. In other words, all it takes is that, contextually, the quoted words be an exemplification of the relevant type.

6. Closed and Open Quotations

In addition to illustrating linguistic demonstrations, example (4) exhibits another important characteristic: the quotation in it functions as a nominal constituent. This nominal constituent is a singular term that refers to a linguistic entity, namely a piece of discourse. But not all quotations behave like that. Not all quotations are *referential NPs*. Consider:

- (5) Gerald said he would “consider running for the Presidency”.
- (6) The girl showed the soldier “her” house in Jaffa.

Using the terminology first put forward by Quine in his *Mathematical Logic*, we can say that the quoted sequence in each of the above is *used* at the same time as it is *mentioned*. It is mentioned because it is envisaged *qua linguistic entity* (as indicated by the quotation marks), and it is used because it plays its ordinary syntactic and semantic role in the embedding sentence. This means, notably, that the sentence does not break down grammatically and semantically if the demonstration is removed.

Now try doing the same with (4). What you get is:

- (4₁) Keith said, I think tha’’s all a pile of roobish.

(4₁) is not just (4) without quotation marks. (4₁) is to be read as no longer including any direct quotation. The fact that this proves a difficult exercise is certainly grist to the mill of those who assume (4) to be essentially different from (5) and (6). If (4₁) is grammatical and interpretable at all, its meaning is only remotely related to that of (4). For instance, in the absence of the quotation, “I” no longer refers to Keith but to the utterer of (4) and the judgment expressed by the *that*-clause must be ascribed to that utterer too.

Recanati calls (4) a *closed* quotation, while (5) and (6) illustrate varieties of *open* quotation.¹² The difference between closed and open quotations is a matter of *linguistic recruitment*: in (4), says Recanati, the demonstration (i.e. the *act* of illustrating by exemplification) is recruited to fulfil the syntactic function of direct object. I cannot insist enough on the idea that

¹¹ For more detail on how Recanati articulates types with properties, see his (2001, p. 640f)

¹² In Recanati’s framework, this term has a wider application than I show here. Thus, “free-standing” direct speech not governed by a reporting verb is another variety subsumed under *open quotation*. In the following, to avoid distorting Recanati’s terminology and causing misunderstandings, I shall use the term *hybrid quotation* to refer to the phenomena illustrated by (5) and (6). (See issue 17 of the *Belgian Journal of Linguistics*).

7. Closed demonstrations

(7) Galileo said, “E pur si muove”.

(8) “Tropstical” doesn’t mean a thing.

(9) And the Martian guy shouted, “fgey cvjrv=\$:/b”ç”cà”=+:/ à)à)à)à))àva”.

Diagram illustrating the syntactic structure of the sentence "Galileo said 'E pur si muove'" (Galileo said 'And yet it moves'):

```
graph TD
    S --> NP1[NP]
    S --> VP[VP]
    NP1 --> N[N]
    N --> Galileo[Galileo]
    VP --> V[V]
    V --> said[said]
    VP --> NP2[NP]
    NP2 --> Dem[Dem]
    Dem -.-> Quote["'E pur si muove'"]
```

The structure shows the sentence (S) branching into a Noun Phrase (NP) and a Verb Phrase (VP). The NP branches into a Noun (N), which is "Galileo". The VP branches into a Verb (V), which is "said", and another NP. This second NP branches into a Determiner (Dem), which is "[Dem]", leading to the quoted phrase "E pur si muove".

At the risk of looking like I am quibbling, the sequence “E pur si muove” is not actually part of the syntactic structure: it is the demonstration that is: [Dem]_{NP} is a terminal node, just like *Galileo* and *said*. “E pur si muove” just happens to be the “vehicle” of the demonstration (the token used to perform the demonstration). I have chosen to represent this link by means of a dotted line. Had I been rigorous, the displayed token should not have been featured in the tree diagram at all.

$$\text{NP} \Rightarrow [\text{Dem}]_{\text{NP}}$$

¹³ The latter insight goes back at least to Davidson (1979, p. 37).

¹⁴ The symbol “[Dem]_{NP}” is Recanati’s notation for a demonstration recruited as an NP. For the purposes of this paper I could have been content with the one symbol *[Dem]*, but the subscripts are certainly reader-friendly.



- (10) John went
- (11) Piano student plays passage in manner μ
 Teacher: It's not [plays passage in manner μ] —it's [plays passage in manner μ'].
 (originally in Horn 1989, p. 564)

Though both examples explicitly involve a non-verbal demonstration, a minor difference can be observed. In (10), we have an ordinary instance of mimicry. The sound produced by John is not a camel belch, though it is meant to be close enough for the intended demonstratum to be identifiable. Most cases of non-linguistic demonstrations occurring in discourse actually tend to be like (10): demonstrations of bird songs, doors closing, water dripping from the ceiling, objects crashing into each other, etc. But some are like (11), which is distinctive in that there is no difference in kind between the displayed token and the demonstratum.¹⁵

In the light of the examples provided in this section, I believe it would be arbitrary to draw a boundary between “genuine” (i.e. entirely verbal) quotations and non-verbal demonstrations (cf. Clark & Gerrig 1990, p. 782; Clark 1996, p. 178). For instance, is the demonstration of a pseudo-word (8) or of some Martian utterance entirely verbal? Is it more deserving of the label *quotation* than the demonstrations in (10) and (11)? Note in passing that these could be replaced by onomatopoeia (i.e. something more word-like) without their import being altered. If the previous point is granted, then the next step to take is to propose an account of (10) and (11) in terms of linguistic recruitment. This is easily done for (11): the demonstration of how to play the piano passage is incorporated into the sentence as an NP. In (10), things are slightly less straightforward. After the reporting verb “to go”, we have an adjunct rather than an NP. In Recanati (2001), all recruited demonstrations are recruited as NPs. Recanati, however, makes allowances for recruitment in other syntactic capacities, notably as a common noun (2001, p. 649, fn). Barbara Abbott has shown that recruitment indeed needed to be extended to at least the Adjective Phrase and the Noun positions (2005, p. 15).¹⁶ In (10), we will say that the demonstration is recruited as an Adjunct to “went”. Here is the relevant part of the tree diagram for (10):

¹⁵ I initially thought that this meant that *all* the aspects of the demonstration were depictive. But that is not necessarily the case, as Neftali Villanueva (p.c.) has pointed out to me: the fingers with which particular notes are played may be depictive aspects or not, the sound volume might be or not (the teacher might be annoyed and play too loud), etc.

¹⁶ See also Clark & Gerrig (1990, p. 771f, 782) – they talk of “embedding” a demonstration. In chapter 6.2.2.2 and 6.2.3.3 of my PhD dissertation, I talk at some length of quotations recruited as Ns (see De Brabanter 2002).

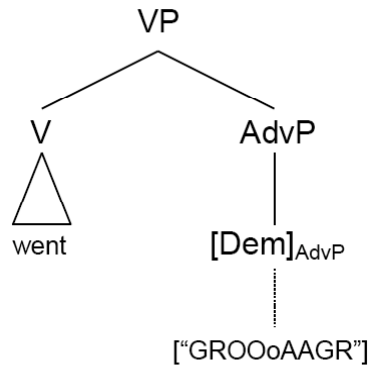


Figure 2.

The same remarks apply here as in the case of (10). The phrase-structure component of the grammar need only be increased by one rule:

$$\text{AdvP} \Rightarrow [\text{Dem}]_{\text{AdvP}}$$

Other recruited demonstrations will require enriching the grammar by a very restricted set of rules like the following:

$$\begin{aligned} \text{AdjP} &\Rightarrow [\text{Dem}]_{\text{AdjP}} \\ \text{PP} &\Rightarrow [\text{Dem}]_{\text{PP}} \\ \text{VP} &\Rightarrow [\text{Dem}]_{\text{VP}} \\ &\dots \end{aligned}$$

It is clear that the recruitment account assumes that demonstrations can only be recruited as constituents: in other words, they should not be found to occupy the same position as strings of words that are *not* constituents. This would necessitate empirical support of a kind that I cannot supply at the moment: first, a corpus of examples should be developed, which is no small undertaking. Second, given the partly non-verbal nature of the examples, it might not be easy to make absolutely sure that a given demonstration stands for a constituent or does not. Space limitations prevent me from going any further into this issue here.

The analysis in figure 2 is supported by the fact that “to go” is used indifferently to report speech or non-linguistic sounds. Take:

(12) And then Sheila went “I think Mike is a better guitarist than Rob”.

It is easy to accept that the complement of “to go” is a direct speech report (as it would be after “to say”). Moreover, there is no good reason to contend that we are dealing with a hybrid quotation, i.e. that the sequence in quotation marks is used at the same time as being quoted: it does no other job than demonstrating (some aspects of) a previous utterance of Sheila’s. Both this unambiguous semantic function and the grammatical similarities with “to say” point to the fact that the quotation must have been recruited syntactically, albeit as an adjunct. There is every reason to extend this analysis to the imitation of the camel belch: this bit of mimicry too is a demonstration, intended as a depiction of a genuine camel belch. Therefore, it must also be a recruited demonstration.

On the theory I advocate, there is no non-verbal counterpart to hybrid quotations. This is not surprising if one bears in mind that the hybrid quotations we have been looking at involve the simultaneous use and mention of a segment and if one further realises that no non-linguistic

segment can ever be *used* linguistically in any meaningful sense. A difficulty therefore arises for the analogy I have drawn between quotations and non-verbal demonstrations. I grant that quotations are not “just demonstrations”, there is something special about them. But what makes them special is precisely the fact that they are *linguistic* demonstrations, a property which automatically endows them with the ability to be *used* in a linguistic environment. I believe that this basically accounts for the difference with non-verbal demonstrations pointed out above. I also believe that the analogy between quotations and non-verbal demonstrations is not threatened by this difference.¹⁷

8. Some more illustrations, some more problems

My corpus of clear-cut examples of recruited non-verbal demonstrations has been very limited so far. Here are a few more examples that may be amenable to the treatment advocated above:

(15) Of course he made a point of looking very



(16) And then he



(17) I got out of the car, and I just [DEMONSTRATION OF TURNING AROUND AND BUMPING HIS HEAD ON AN INVISIBLE TELEPHONE POLE]. (from Clark & Gerrig 1990, p. 782)¹⁸

It is easy to see how (15) can be lumped together with (1), (10) and (11): the demonstration does the same job that an adjective (f.i. “serious” or “glum”) would do, and it does so within a recognisable sentential framework. One is tempted to apply the same analysis to (16) and (17), but there is an additional difficulty: there is no clear sentential pattern here, so that it is less safe to claim that the structures in (16) and (17) are headed by an inflectional element.

However, we are still dealing with very different examples from Stainton’s. All of his examples can, at least at first sight, be identified as structures headed by a lexical category, which is not the case here: in (16) and (17), the structure is “waiting for a head”, and this head, which is inflectional, is only provided by the non-linguistic demonstration. Moreover, Carston’s “on the shelf” and most of Stainton’s examples are self-sufficient in a way that (16) and (17) are not: “Nice dress”, “Dinner for seven”, “From Brazil”, “Chunks of strawberries” are all complete phrases (NPs, PPs) which, given a context of utterance, can be produced to perform a speech act. In (16) and (17), however, we have structures that are incomplete and cannot serve to perform a communicative act unless they are supplemented with an act of non-verbal communication. Therefore, we are not dealing with Stainton-like cases. Still, I admit that (16) and (17) would be amenable to an analysis that is different from the one I have

¹⁷ There is, though, a difference that deserves to be investigated: whereas quotations are in essence metarepresentational, the iconic gestures in my data do not represent representations but “world-level” properties, actions, or individuals. Of course, I could gesture echoically to represent someone else’s communicative gesturing, but this is by no means necessary.

¹⁸ Levinson (2004, p. 101) has two examples of utterances in which an indexical gesture “replaces a linguistic expression”. More can be found in Clark & Gerrig (1990, p. 781f).

suggested: it could be said that they are not utterances of sentences with a recruited demonstration but rather just multi-modal messages combining verbal and non-verbal stimuli. Note that the latter analysis cannot fail, and a further drawback may be that it does not throw any light on the similarities between, on the one hand, (16)-(17) and, on the other, (1)-(10)-(11)-(15) together with the clearly quotational cases. It would also be useful to see if the opposite case to (16)-(17) can be observed too: an “utterance” begins as a piece of non-verbal communication and ends with a recognisable verbal structure. Would such cases also be analysed as suggested in terms of recruitment, or would we have to grant that we are “just” dealing with multi-modal messages deprived of a genuine linguistic structure?

9. Ellipsis revisited

In this final section, I want to show that the recruitment-based account does not place me in the same camp as Stanley, Ludlow or Merchant. In other words, it is not because I have not adopted Stainton’s analysis to my small corpus of examples that I end up with the opponents of non-sentential speech acts.

First, like Stainton’s, I regard my examples as genuine speech acts, with a recognisable propositional content and illocutionary force. Second, there is every reason *not* to account for them in terms of ellipsis.¹⁹ This point is perhaps less self-evident, so I will devote a few lines to providing some support for it. I begin with my central cases, i.e. cases where a sentential structure is identifiable (or so I argue). It is tempting here to say that some syntactic constituent is missing but can be recovered. For example, in (1), it could be argued that the AdjP “frighteningly grumpy” has been ellipted. The first problem with this suggestion is that several other AdjPs could be supplied with equal success, provided they expressed a sufficiently similar meaning: “very grumpy”, “terrifyingly ill-tempered”, etc. Such freedom is not standard in ordinary cases of ellipsis, where a *precise* constituent is usually retrievable (see Elugardo & Stainton 2004 for a similar point). However, a putative proponent of the ellipsis account could retort that the demonstration raised a certain word or phrase to salience, and that this provided a linguistic antecedent which made syntactic ellipsis possible. Be that as it may, there is a stronger argument: on the analysis advocated here, the claim is that *nothing is missing*! We have a full linguistic structure because a non-verbal communicative act is part of the ostensive stimulus and therefore syntactically recruitable. Once recruited, it occupies a position that *could* have been (but is not) occupied by a linguistic constituent, so there is no place left to be occupied by an ellipted segment!

As regards my less central examples, like (16) and (17), they are clearly not elliptical: too much would have to be conjured out of the magician’s hat for an ellipsis account to be tenable. Still, someone who frontally opposed the idea of syntactic recruitment could argue à la Stanley that they cannot be used to perform a genuine speech act. I am not sure I can counter this objection. And I am not sure I must absolutely. After all, (16) and (17) are borderline cases.

10. Some tentative conclusions

In this paper, I have outlined an extension of Recanati’s theory of quotation. In particular, I have exploited his decisive distinction between open (hybrid) and closed cases, which rests on the validity of the notion of linguistic recruitment. The notion does an excellent job when

¹⁹ I must admit that I have no idea if ellipsis theorists would at all be tempted to account for my examples in terms of ellipsis. But, lest there be misunderstandings, I prefer to explicate the reasons why I think an account in terms of ellipsis would not be satisfactory either.

applied to closed quotations. I think it does too when applied to non-verbal demonstrations. At any rate, it does much better than possible rival accounts in terms of non-sentential speech acts or in terms of syntactic ellipsis.

Yet there remains one major difficulty: there is no obvious cut-off point between the cases that should be accounted for in terms of recruitment and those that should not. I have drawn on a fairly intuitive notion of linguistic structure: if a given utterance displays enough linguistic structure, then many non-linguistic “things” can be incorporated as part of that structure. But how much is “enough”? At this stage, I do not really know. I have to be content with saying that there must be at least a VP (or IP). But that may not be saying much.

In spite of this patent weakness in the account I am proposing, I do believe that my extension of recruitment rests on some strong foundations. Its starting point seems to me to be sound: there is no principled difference between fully verbal closed quotations and “non-verbal closed quotations” (cf. section 7). Therefore, one cannot but admit that the demonstrations in (10) and (11) are recruited. Now, the fact is that these non-verbal demonstrations (which are analogues of closed quotations) do not function differently from those that occur in positions where one might expect *hybrid* quotations, as in (1) and (15). The reason is that hybrid quotations do not have a non-verbal counterpart: gesturing cannot be *used* in the Quinean sense, simply because it is not linguistic. Therefore, it seems entirely justified to account for (1) and (15) too in terms of linguistic recruitment.

Problems arise when we turn to cases like (16) and (17), which display less syntactic structure. It is tempting to analyse examples like these *not as sentences*, just as multi-modal *utterances* or *stimuli*. This, for instance, is the approach that Dan Sperber (p. c.) would favour. Sperber’s position has the advantage of being perfectly straightforward: if you want to avoid giving precedence to linguistic aspects of multi-modal signals, it is safest not to attribute a linguistic role to non-verbal stimuli.

Still, although I wholeheartedly endorse the view that the study of human communicative acts should not be linguistically biased, I think the all-out non-linguistic approach has at least one undesirable consequence: it has nothing to say about the connection between non-verbal demonstrations and quotations. Yet, this connection exists. We are back to square one: (some) non-verbal demonstrations just *are* quotations. How do we account for that without appealing to recruitment?

References

- Abbott, B. (2005) ‘Some notes on quotation’. *Belgian Journal of Linguistics* 17: 13-26.
- Carston, R. (1999) ‘Review of Herbert Clark, *Using Language*’. *Journal of Linguistics* 35: 167-71.
- Carston, R. (2002) ‘Linguistic meaning, communicated meaning and cognitive pragmatics’. *Mind & Language* 17: 127-48.
- Carston, R. (2004) ‘Explicature and semantics’. In S. Davis & B. Gillon (eds.), *Semantics: A Reader*. Oxford: OUP, pp. 817-45.
- Carston, R. & Powell, G. (Forthcoming) ‘Relevance theory – new directions and developments’. In E. Lepore & B. Smith (eds.), *The Oxford Handbook of Philosophy of Language*. O.U.P. Online at www.phon.ucl.ac.uk/home/robyn/home.htm.
- Clark, H. H. (1996) *Using Language*. Cambridge University Press.
- Clark, H. H. (1997) ‘Dogmas of understanding’. *Discourse Processes* 23: 567-98.
- Clark, H. H. (2003) ‘Pointing and placing’. In S. Kita (ed.), *Pointing. Where Language, Culture, and Cognition Meet*. Hillsdale, New Jersey: Erlbaum, pp. 243-68.
- Clark, H. H. & Gerrig, R. J. (1990) ‘Quotations as Demonstrations’. *Language* 66: 764-805.
- Davidson, D. (1979) ‘Quotation’. *Theory and Decision* 11: 27-40.

- De Brabanter, P. (2002) *Making Sense of Mention, Quotation, and Autonymy. A Semantic and Pragmatic Survey of Metalinguistic Discourse*. Unpublished Ms.
- De Brabanter, P. (2005) 'Quotations and the intrusion of non-linguistic communication into utterances'. In A. Dey et al. (eds.), *CONTEXT'05. Lecture Notes in Artificial Intelligence*, 3554. Berlin/Heidelberg: Springer Verlag, pp. 126-39.
- Elugardo, R. & Stainton, R. (2004) 'Shorthand, syntactic ellipsis, and the pragmatic determinants of what is said'. *Mind & Language* 19: 442-71.
- Horn, L. R. (1989) *A Natural History of Negation*. Chicago, London: The University of Chicago Press.
- Levinson, S. C. (2004) 'Deixis'. In L. Horn & G. Ward (eds.), *Handbook of Pragmatics*. Blackwell publishing, pp. 97-121.
- Ludlow, P. (2005) 'A note on alleged cases of nonsentential assertion'. In R. Elugardo & R. Stainton (eds.), *Ellipsis and Nonsentential Speech*. Dordrecht: Springer, pp. 95-108.
- Melinger, A. & Levelt W. J. M. (2004) 'Gesture and the communicative intention of the speaker'. *Gesture* 4: 119-41.
- Merchant, J. (2004) 'Fragments and ellipsis'. *Linguistics & Philosophy*, 27: 661-738.
- Recanati, F. (2000) *Oratio Obliqua, Oratio Recta : An Essay on Metarepresentation*. Cambridge, Mass.: MIT Press, Bradford Books.
- Recanati, F. (2001) 'Open Quotation'. *Mind* 110: 637-87.
- Sperber, D. & Wilson, D. (1995) *Relevance. Communication and Cognition*, Oxford/Cambridge, Mass.: Blackwell.
- Stanley, J. (2000) 'Context and logical form'. *Linguistics and Philosophy* 23: 391-434.
- Stainton, R. (1995) 'Non-sentential assertions and semantic ellipsis'. *Linguistics and Philosophy* 18: 281-96.
- Stainton, R. (2005) 'In defense of non-sentential assertion'. In Z. G. Szabó (ed.), *Semantics versus Pragmatics*. Oxford: Clarendon Press, pp. 383-457.
- Wilson, D. (2000) 'Metarepresentation in linguistic communication'. In D. Sperber (ed.), *Metarepresentations. A multidisciplinary perspective*. Oxford: O.U.P, pp. 411-48.